DOCUMENT RESUME

ED 318 787 TM 014 902

AUTHOR Anderson, Paul S.; Saliba, Alcyone
TITLE Changing World Patterns of Machine-Scored Objective Testing: The Expected Impact of the Multi-Digit Method.
PUB DATE 87
NOTE 7p.; Paper presented at the World Conference of Comparative Education (6th, Rio de Janeiro, Brazil, July 6-10, 1987).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Computer Assisted Testing; Developed Nations; Developing Nations; *Educational Technology; Educational Trends; Foreign Countries; Higher Education; *International Studies; *Objective Tests; Test Format; Testing Problems; *Test Scoring Machines; Trend Analysis
IDENTIFIERS *Multi Digit Tests

ABSTRACT

The use of optical scanners and computers in educational testing is common where objective testing methods (such as true-false, matching, and multiple-choice items) are well-established means of evaluating educational achievement. Where non-objective testing methods (such as fill-in-the-blank, short-answer, and essay items) have been more common, however, the diffusion of automated test scoring processes may be slow. A classification model of world patterns of educational testing methods at the university level is outlined. The three patterns are characterized as: (1) maximum current usage of machine scoring, as found in introductory courses with large numbers of students; (2) little use of machine scoring in spite of the financial resources to do so; and (3) very little use of machine scoring and few financial resources to support it. The first pattern mainly signifies universities in the United States; the second model refers to universities found throughout Europe and some American schools; and the third pattern refers mainly to developing countries in Latin America, Africa, and Asia. The ways in which these patterns are expected to change in response to contemporary demographic, economic, and technological factors are discussed. One technological factor is the recent development of the Multi-Digit Testing technique, which generates computer-scorable test items equivalent to fill-in-the-blank items, thus combining the academic rigor of free recall items with up-to-date educational testing technology. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED318787

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
PAUL S. ANDERSON
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CHANGING WORLD PATTERNS OF MACHINE-SCORED OBJECTIVE TESTING: THE EXPECTED IMPACT OF THE MULTI-DIGIT METHOD

Paul S. Anderson
Illinois State University

Alcyone Saliba
Illinois State University

TM014902

BEST COPY AVAILABLE

Changing World Patterns of Machine-Scored Objective Testing: The Expected Impact of the Multi-Digit Method.

Paul S. Anderson
Dept. of Geography-Geology
Illinois State University
Normal, IL 61761

Alcyone Saliba
Dept. of Educational Administration
and Foundations
Illinois State University
Normal, IL 61761

(A paper presented to the Sixth World Congress of Comparative Education, Rio de Janeiro, Brazil, 6-10 July, 1987.)

ABSTRACT

The use of optical scanners and computers in educational testing is common where objective testing methods (such as those including true/false, matching, and multiple-choice items) are well established means of evaluating educational achievement. Non-objective testing methods (such as those including fill-in-the-blank, short-answer, and essay items), on the other hand, have not promoted the diffusion of automated test scoring processes. This paper (1) outlines a classification model of world patterns of educational testing methods and (2) discusses how these patterns are expected to change in response to contemporary demographic, economic, and technological factors. One technological factor is the recent development of the Multi-Digit Testing (MDT) technique, which generates computer-scorable test items equivalent to fill-in-the-blank ones, thus combining the academic rigor of free recall with up-to-date educational testing technology.

I. INTRODUCTION

The use of optical scanners and computers to score educational tests and exercises is well established in some international regions but not in others. Where implemented, machine scoring has resulted in substantial usage of three formats of "objective" educational testing: true/false, matching, and multiple choice. The regions of the world emphasizing the use of written responses (namely fill-in-the-blank, short answer, and essay questions) have not widely adopted machine scoring. The latter regions can be divided into countries with sufficient financial resources to acquire optical scanners and those without such resources. The use of machine scoring in these three divisions (models)' is expected to change in response to demographic, economic and technological factors. The current situations are described in three models. Reasons for and expected results in the shifts in the patterns are then discussed.

II. THREE INTERNATIONAL MODELS

The focus of these basic models is on university level education. Parallels in secondary education exist but are not emphasized in this paper. The basis of the paper is mainly from observations and experiences by the authors on three continents. An ongoing search for relevant published references has yielded minimal information on international usage of test formats.

TM014902

A. The "A" model represents the maximum current usage of machine scoring in education. This model is typified by introductory college courses with large enrollments of 50 to 500 students. More precisely, the student-to-faculty ratio is high. The "A" could almost signify "America" (USA) except that small enrollment, prestigious and expensive American universities which focus on small classes are not in this model.

The USA does provide the best examples of this first model. A nationwide but not randomly selected survey of seventy-one university testing and measurement offices (Erwin, Chatman and Nelson, 1985) obtained the following information. With an average enrollment of twenty-one thousand students, those universities scan an average of 481,000 documents (pages) per year. Of those, about half are answer sheets from classroom tests (authors' estimate based on personal interviews). The other half includes research surveys and administrative work such as teaching evaluations and counseling studies. The machine-scored classroom tests therefore average approximately ten per student per year. Since few of those tests would be for upper classmen (juniors and seniors) and graduate students, the average is probably closer to twenty machine-scored tests per year per freshman or sophomore. If small-class courses such as English, foreign languages and speech is excluded, there are approximately three machine-scored tests per semester course. It is not uncommon in large-enrollment classes that 100 percent of all testing is by machine-scored objective methods. Apart from speed of grading, one major advantage enjoyed at these universities is the quantity and quality of computer-generated feedback.

A less intensive variation of Model "A" is found in smaller schools where other priorities for financial resources prevent the offering of full-service machine scoring. Although varying greatly between professors, the attitude toward objective tests is one of general acceptance. True/False, matching, and multiple choice questions are often used with manual scoring. Sometimes a hole-punched template key is used. Other times the response letters of A through E are recorded down the left hand margin of the page. Alternatively, some schools have old-fashioned scanners that do not connect to any computer. Once again, class size per instructor is a major factor in the use of objective tests. Given that over half of all American youths enter some form of formal post-secondary education, it is easy to see why classes can be so large.

Readers should <u>not</u> have the impression that all of American education uses only objective tests for classroom assessment. Written term papers and essays are commonly requred. Also, numerous universitites are very much like those in Model "B" described below.

B. Universities in the second model, "B", have the financial resources to acquire machine scoring and item banks, but they seldom, if ever, utilize "objective" test methods. These schools are found throughout Europe and in those select American schools not included in the "A" model. Class sizes are small and the ratio of students to instructors is low. The British educational tradition provides good examples.

Part of the British university tradition is embodied in the Oxford and Cambridge systems of tutorialized education. The idealized personal contact of student with professor can be traced back to ancient Greek civilization. Although a desirable method to teach a highly select and quite small body of students, the tutorialized method is not widespread because of costs.

Examples of mainstream British university education was as observed by one author in Australia for six years (1972-1977). No use of objective test methods was observed in any testing or exercises. Essay questions dominated. A few short answer responses were used. Furthermore, the British system favors infrequent evaluation by test. Comprehensive tests at the end of three years of study were formerly the norm. Now, end of year and end of semester testing are typical. Tests and exercises for credit, (that is, "continual assessment") during the semester is uncommon and seldom more than thirty precent of the total evaluation of a course.

Although not common, large classes do exist in introductory courses. The largest one observed was "Introduction to Geography" at the University of New England, Armidale, NSW. Approximately two hundred internal (in attendance) and three hundred external (distance education) students were enrolled. Six faculty members divided the lecturing. Seven half-time teaching fellows (graduate assistants) were responsible for the tutorials of 7-12 students and the laboratory practical exercises. The latter could have well utilized some objective techniques. But with substantial staff resources available plus a strong tradition favoring essay and short answer responses, the objective methods were not considered to be a viable alternative.

In the British sphere of university education, the one distinct move toward partial usage of machine scoring came with the development of the Open University (OU) in the United Kingdom. The OU was specifically established by politicians to make university education more accessible to larger numbers of students. Large classes with thousands of students were part of the original plan. Machine scoring of objective tests and exercises has been one means used to help fulfill the OU mandate.

C. Model "C" refers mainly to the developing countries in Latin America, Africa and Asia. Except for entrance examinations, such as the Brazilian "vestibular," there is very little usage of machine-scored testing in Third World countries. Manual scoring of objective tests is also uncommon. Most of these countries have educational systems based on those of their former European colonial governments. US American influence in education is present, but relatively recent. The American model with objective tests and machine scoring is certainly well known to educational leaders in the developing countries. However, primarily because of tradition plus financial reasons reflecting the cost of hardware (optical scanners and computers), the use of machine scoring and objective tests is quite low. Furthermore, class sizes are relatively small. Although population pressures and the need for skilled graduates are great and increasing, relatively few students can find university places in most developing countries.

Even at the relatively new, relatively progressive and relatively well supported University of Brasilia, the typical class size is only twenty students. The authors studied and lectured there for over four years (1978-1982). Several large (approximately 250 seat) lecture theaters are in the "Minhocão," but they appeared to be seldom used, seldom filled and not liked by either students or faculty.

III. FACTORS TO AFFECT CHANGE

The usage of machine scoring is expected to change dramatically but not uniformly in these three models. There are three key factors: demography, economics, and educational technology.

1) Demography: Growing populations and/or changing expectations of people are major formentors of change. Governments need to respond to those pressures by expanding the educational programs. 2) Economics: The costs of optical scanners and microcomputers are dropping rapidly. The financial reasons to not use machine scoring will soon disappear. 3) Educational technology: Recent developments will broaden the potential and impact of machine-scored testing and exercises.

One technological factor is the recent development of a computer-scorable near-equivalent to fill-in-the-blank questions. That method, called the MDT Multi-Digit Testing Technique, offers the increased academic rigor of free recall (versus recognition) of correct responses. The MDT method is described in the recent book by Anderson (1987). "This book discusses modern, readily available technology that can improve education through machine scoring of tests and exercises. The MDT innovation is not just a book, it is an accomplished reality in computer software and support materials. Stripped to its fundamentals, the testing method requires students to know (not just recognize) their answers and to obtain the corresponding label numbers from an available list of responses. Using a short list, the MDT multi-digit technique is similar to matching. Using a long list with hundreds of alphabetized responses, any test becomes a close approximation to a fill-in-the-blank exam. The label numbers are marked by pencil on answer sheets that can be read manually or by a machine. There are numerous advantages which result from this applied innovation, including cost savings for schools, time savings for teachers, enhanced feedback via computer analysis, and, most important, improved learning for students." The knowing of factual information is one important foundation of education, especially in technical and professional fields like medicine and engineering.

IV. EXPECTED CHANGES

In the "A" model, the MDT method is expected to result in a reduction in the number of multiple choice questions asked in American educational testing. It is estimated that 40 percent of current multiple choice questions could be easily transformed into the more rigorous MDT multi-digit format. Furthermore, because the MDT method is a machine-scorable near-equivalent to fill-in-the-blank questions, an even greater percentage of all testing in the USA will shift to the new expanded definition of machine scorable. In essence, the fill-in-the-blank segment of the previously manually scored techniques can now be incorporated into the realm of machine scoring. Especially when numerical responses to mathematical problems are calculated, the MDT technique offers a superior way to record student answers for machine scoring. The increase in usage plus the lowering of the prices of the machine scoring hardware should result in a much greater sale of such equipment in the Model "A" environment where objective tests and machine scoring are already deemed acceptable.

The MDT method overcomes the frequent complaint that students are able to either recognize, select by elimination, or outright "guess" correct responses as found in common multiple choice questions. For Model "B" situations, the MDT method is expected to attain wider acceptance than has the multiple choice method. The MDT technique will allow European instructors to more rapidly, more thoroughly and more frequently assess the factual or discrete-answer knowledge of their students. This will provide the instructors with more time to devote to research and to issues of higher order learning, including grading essays or conducting tutorials. Nevertheless, the forces of traditionalism are very strong.

Increased usage of machine scoring in Europe will most likely occur where administrative and political influence, such as the creation of the Open University, stimulates the offering of educational opportunities to larger segments of the population.

In Third World countries, (Model "C"), the major impact is expected to come from the lowering of prices of all components (hardware and software) to less than US$1500 within a few years. The first impact will occur at the universities that decide to respond to the pressing need to teach many more students. Considering the population pressures and the national needs for trained workers in developing countries, machine scoring with the rigorous MDT multi-digit method should grow dramatically and yield positive results.

V. CONCLUSION

Although machine scoring of objective tests cannot by itself resolve the world education crisis, it is fast becoming an economical and academically powerful tool to reach large numbers of students. The future of educational testing is highly likely to include the expansion of capabilities and the increased usage of machine scoring world wide.

BIBLIOGRAPHY

Anderson, Paul S. (1987) The MDT Innovation: Machine-Scoring of Fill-in-The-Blank tests. Multi-Digit Technologies Corporation (P.O. Box 14), Normal, Illinois, 61761 USA

Erwin, T. Dary, Steven Chatman and Randall B. Nelson, (1985) "Survey of Characteristics of Measurement Service Offices," Conference paper to The Measurement Services Association of the National Council on Measurement in Education. c/o Texas A & M University.

Helmick, John S. and Sanford C. Jameson (1977) "Examinations and Tests," International Encyclopedia of Higher Education, vol. 9, pp. 1498-1505.